

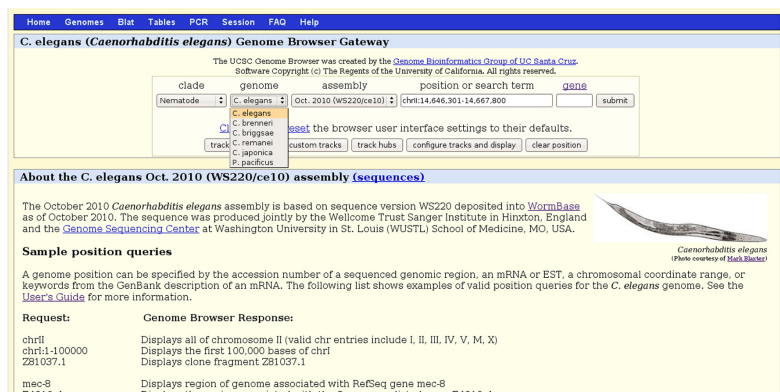
# An introduction to the UCSC Genome Browser\*

Raymond Lee<sup>§</sup>

Division of Biology, California Institute of Technology, Pasadena CA 91125, USA

The [Genome Browser](#) at the University of California Santa Cruz provides a uniform graphical interface to sequences, features, and annotations of genomes across a wide spectrum of organisms, from yeast to humans. In particular, it covers seven nematode genomes: *Caenorhabditis elegans*, *C. sp. 11*, *C. brenneri*, *C. briggsae*, *C. remanei*, *C. japonica*, and *Pristionchus pacificus* and thus is particularly useful for multiple-genome comparative analysis. One can use the provided tools and visual aides to facilitate sequence feature detection and examination. This article provides a brief introduction for using the [Genome Browser](#) from the perspective of a *C. elegans* researcher. Interested users should read the official user guide and explore the site more deeply as there are many more features not mentioned here.

To begin, one should choose the nematode clade, then the organism *C. elegans*, a version of genome sequence assembly (e.g., WS220), and a region of the genome ([Figure 1](#)). Browser behavior is context sensitive, depending on the choices made in order from left to right. There are three basic ways to specify a region to browse: by a range of chromosomal numerical positions, by a gene name and by an accession number. More flexibly, instead of exact positions, one can also search for a descriptive term (such as “kinase inhibitor”) that is present in gene records. Some restrictions are worth noting. Chromosome positions must start with “chr”. Gene names recognized by the browser are limited to those in RefSeq and Ensembl entries. Remarkably, most if not all named miRNA genes, like *lin-4*, *let-7*, *lxy-6* and 170 *mir*- loci are excluded from the [Genome Browser](#).



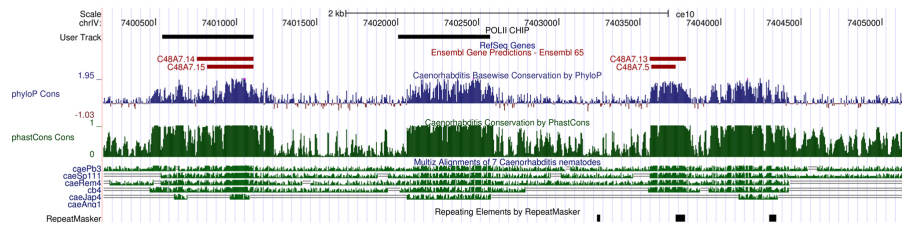
**Figure 1. UCSC Genome Browser select a region.** A screen capture of the UCSC Genome Browser that illustrates the process of selecting a specific region in the *C. elegans* genome to browse.

\*Edited by Oliver Hobert. Last revised April 23, 2013, Published July 11, 2013. This chapter should be cited as: Lee R. An introduction to the UCSC Genome Browser (July 11, 2013), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.160.1, <http://www.wormbook.org>.

**Copyright:** © 2013 Raymond Lee. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

<sup>§</sup>To whom correspondence should be addressed. Email: [raymond@caltech.edu](mailto:raymond@caltech.edu)

Once a region is selected, the **Genome Browser** can display a wide selection of sequences and features as tracks. The tracks are easily customizable. One can add or remove specific tracks and change the order or look of tracks to make them easier to browse. Several tracks based on multiple genomic sequences are grouped under the Comparative Genomics section. There is a multiple sequence alignment track (Multiz) and three possible conservation tracks (phyloP, phastCons and conserved elements) to graphically display the results of conservation analyses. Scanning the genome with these tracks turned on (Multiz and phyloP in the following example), one can quite easily spot areas that show significant cross-species conservation, possibly implicating functional domains. For instance, on chromosome IV, between C48A7.14 and C48A7.13, there is a 530 bp region (chrIV:7402060-7402590) that is highly conserved between *C. elegans* and five other species but there is no existing annotation (Figure 2). It could be an extension of a gene model nearby or a new gene.



**Figure 2. Genome Browser alignment of diverse sequence features.** A Genome Browser view of a 4.5 kbp genomic region of *C. elegans* LGIV. From top to bottom, there are aligned tracks of user-defined features (see text for explanation), gene models, three types of multiple sequence comparisons, and repeat elements.

One important extension of the **Genome Browser** is the ability to add custom tracks. Custom tracks allow users to incorporate their own or third party data into the browser so that imported sequence annotations can be inspected in the context of existing ones. A large collection of diverse sequence-based experimental data have been published by the **ModENCODE** project. There are experiments assessing every part of the genome for its involvement in expression regulation under specific biological conditions. **ModENCODE** data sets may be downloaded as sequence features in a compatible GFF3 format and included as custom tracks in the **Genome Browser**. As an example, I added a track for `modencode_174` (labeled as “User Track:POLII CHIP”), which marks genomic regions that have been identified as potential binding sites for RNA polymerase II. There is substantial overlap between a potential Pol II binding site and the inter-species conserved sequence between C48A7.14 and C48A7.13. This observation further supports the possibility that this stretch of sequence is functionally important and may be expressed.



All WormBook content, except where otherwise noted, is licensed under a **Creative Commons Attribution License**.